JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 60, 477–482 (1977)

On Model Reduction of Large-Scale Systems

T. T. SOONG

Faculty of Engineering and Applied Sciences, State University of New York at Buffalo, Buffalo, New York 14214

1. INTRODUCTION

In dynamical studies of large-scale systems, aggregation techniques enable one to represent a system dynamically by a simpler system so that the desired output of the original system can be approximated by that associated with the simpler model. These techniques are becoming increasingly attractive in the study of engineering, urban, environmental, and other systems possessing high dimensionality and not easily accessible to state measurements. The generation of simple models yields distinct advantages from computational as well as modeling points of view. The computational requirements in system simulation, for example, are reduced by a factor equal to the ratio of the orders of the reduced and original systems. Even more savings are realized in control calculations since the length of computation for system control varies with the square of the system order.

The model reduction problem for time-invariant linear systems can be stated as follows: Consider an nth-order linear system S_1 defined by

$$S_1: \qquad \dot{x} = Ax + Bu \tag{1}$$

where x is the n-dimensional state vector, A is an $n \times n$ system matrix, and u is a p-dimensional input vector. Let z be an m-vector ($m < n$) related to x by

$$z = Cx. \tag{2}$$

It is desired to find an mth-order system S_2 described by

$$S_2: \qquad \dot{z} = Fz + Gu. \tag{3}$$

The $m \times n$ matrix C in Eq. (2) is the aggregation matrix and S_2 is the aggregated system or the reduced model. The matrices F and G in Eq. (3) are to be determined.

ISSN 0022-247X

In determining these matrics, it is easy to verify that F and G satisfy

$$FC = CA, \qquad G = CB \tag{4}$$

If C is specified, the matrix G is directly determined from the second relation. The matrix F, however, requires approximation since the first relation constitutes an overspecified system of equations.

Several methods of approximation now exist. One approach is to consider a reduced-order system which contains the most dominant eigenvalues of the original system. Methods proposed by Davison [1], Chidambara [2], and Mitra [3] are based upon this approach. However, these procedures generally require the reduction of the state transition matrix into canonical form. This task becomes cumbersome if some of the eigenvalues are complex or repeated.

Alternate methods include those discussed by Anderson [4] and Fellows *et al.* [5] who select the coefficients of the reduced model in such a way that the corresponding responses of the original and reduced systems are approximately matched. The determination of these coefficients becomes time-consuming if the order of the reduced system is not sufficiently small.

The simplest procedure is an ad hoc one proposed by Aoki [6] who approximates F by

$$F = CAC^T(CC^T)^{-1}, \tag{5}$$

where T and -1 denote matrix transpose and matrix inverse, respectively. The rank of C is assumed to be m.

It is, however, diffcult to assess the aggregation error. The error $e(t)$ as defined by

$$e = z - Cx \tag{6}$$

is governed by the differential equation

$$\dot{e} = Fz - CAx = Fe + (FC - CA)\, x \tag{7}$$

whose solution is

$$e(t) = \exp(Ft)\, e(0) + \int_0^t \exp[F(t-s)]\,(FC - CA)\, x(s)\, ds. \tag{8}$$

It is seen that the evaluation of this aggregation error requires the knowledge of $x(t)$, the solution of the original system. This runs counter to the motivation behind using reduced models.

In what follows, a different approach is used for approximating the matrix F. A significant result produced using this approach is a simple expression which can be used for aggregation error assessment.

2. A Statistical-Estimation Approach

An alternate way to determine an approximation of F is to recast the problem into one in statistical estimation. We rewrite the first relation in Eq. (4) as

$$CA = FC + E, \tag{9}$$

where E is the "error matrix." This equation defines a multivariate linear regression model in which F is to be estimated from the "observation matrix" CA. An optimal estimate of F from this point of view can be used as a good approximation for F and, more important, the covariance matrix associated with the estimate, which is independent of the solution of the original system, can be used to assess the aggregation error.

These results are stated below as a theorem. First, some definitions and known results needed for this development are stated.

Definitions. (a) The Kronecker product of two matrices A and B is defined as [7]

$$A \otimes B = [a_{ij}B]. \tag{10}$$

If A is $n \times m$ and B is $p \times q$, then $A \otimes B$ is $np \times mq$.

(b) Let the "pack" operator be denoted by "vec." Then vec A is a vector formed by stacking of columns of A from left to right into a single vector, i.e.,

$$\operatorname{vec} A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix}, \tag{11}$$

where A_j is the jth column of A.

Listed below are some matrix and vector identities useful for our development. They can be found in Neudecker [8, 9]. Assuming that all indicated operations exist, we have

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD), \tag{12}$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}, \qquad (A \otimes B)^T = A^T \otimes B^T, \tag{13}$$

$$\operatorname{vec} AB = (I_p \otimes A) \operatorname{vec} B = (B^T \otimes I_n) \operatorname{vec} A. \tag{14}$$

In Eq. (14), A is $n \times m$ and B is $m \times p$ and I_j represents a $j \times j$ identity matrix.

We also cite well-known results in multiple linear regression. Consider

$$y = A\beta + n \tag{15}$$

relating the n-vector y of dependent variables, the $n \times p$ matrix A of explanatory variables, and the statistically independent components of n, each having mean zero and variance σ^2. The linear, unbiased, minimum variance estimate of β, denoted by $\hat{\beta}$, is

$$\hat{\beta} = (A^T A)^{-1} A^T y \tag{16}$$

with covariance matrix

$$\operatorname{cov}(\hat{\beta}) = \sigma^2 (A^T A)^{-1}. \tag{17}$$

THEOREM. *In the model of multivariate linear regression*

$$CA = FC + E, \tag{18}$$

where the matrix CA of dependent variables is $m \times n$, the matrix C of explanatory variables is $m \times n$, and the elements of E are statistically independent with zero means and identical variances σ^2, the linear, unbiased, minimum-variance estimate of F, $\hat{F}$, is

$$\hat{F} = CAC^T(CC^T)^{-1} \tag{19}$$

and the covariance of $\hat{F}$ is

$$\operatorname{cov}(\operatorname{vec} \hat{F}) = \sigma^2[(CC^T)^{-1} \otimes I_m]. \tag{20}$$

Proof. Performing "vec" operation on Eq. (18) and using Eq. (14) gives

$$\begin{aligned} \operatorname{vec}(CA) &= \operatorname{vec}(FC) + \operatorname{vec} E, \\ &= (C^T \otimes I_m) \operatorname{vec} F + \operatorname{vec} E. \end{aligned} \tag{21}$$

This is in the form (15) of multiple linear regression. Hence, in accordance with Eq. (16), we have

$$\operatorname{vec} \hat{F} = [(C^T \otimes I_m)^T (C^T \otimes I_m)]^{-1} (C^T \otimes I_m)^T \operatorname{vec}(CA). \tag{22}$$

In view of identities (12)–(14), we have

$$\operatorname{vec} \hat{F} = \operatorname{vec}[CAC^T(CC^T)^{-1}]$$

and we obtain Eq. (19). Similarly, we see from Eq. (17) that

$$\operatorname{cov}(\operatorname{vec} \hat{F}) = \sigma^2[(C^T \otimes I_m)^T (C^T \otimes I_m)]^{-1},$$

which reduces to Eq. (20) upon using identities (12), (13).

It is noted that the result for $\hat{F}$ is identical to that of Aoki. An additional result produced by this approach is the covariance of $\hat{F}$ as given in Eq. (20). It can be used in assessing the accuracy of using Eq. (19) as an approximation of F. Since some freedom always exists in the choice of the aggregation matrix C, one can choose an appropriate C in such a way that the covariance matrix defined by Eq. (20) is at some acceptable level. For example, the trace of the covariance matrix may be minimized with respect to an admissible set of C matrices. We note that minimizing the trace of Eq. (20) is equivalent to minimizing the trace of $(CC^T)^{-1}$.

Another possible application of this result is in assessing the trade-off between aggregation accuracy and the order of the aggregated system. When C is specified, it is sometimes of interest to ask whether aggregation accuracy can be significantly improved by adding another row to the C matrix or, equivalently, by increasing the order of the reduced system by 1. Let

$$C' = \left[\frac{C}{C_1}\right],$$

where C_1 is a row matrix. Optimal choice of C_1 can be first determined by minimizing the trace of the associated covariance matrix. The covariances of $\hat{F}$'s corresponding to C and C' can then be compared to see whether a significant improvement is possible by adding C_1 to the aggregation matrix. This procedure was fruitfully applied to an aggregation problem in transportation [10].

Finally, let us remark that the main results in Eqs. (19) and (20) are obtained under a simple covariance assumption on the error matrix E. There are situations in which modeling of the original system leads to a system matrix A whose elements are specified with different levels of confidence. Since A appears in the dependent variables in our multivariate model, this can be properly taken into account by assuming an appropriate covariance structure for E. A different estimate of F and its associated covariance would result, further demonstrating the utility of this approach.

References

1. E. J. Davison, A method for simplifying linear dynamic systems, *IEEE Trans. Automatic Control* **AC-11** (1966), 93–101.
2. M. R. Chidambara, Two simple techniques for the simplification of large dynamic systems, *in* "Proc. 1969 JACC," pp. 669–674, 1969.
3. D. Mitra, The reduction of complexity of linear, time-invariant dynamical systems, *in* "Proc. 4th IFAC," pp. 19–33, 1969.
4. J. H. Anderson, Geometrical approach to the reduction of dynamical systems, *Proc. IEE* **114** (1967), 1914–1918.
5. D. M. Fellows, N. K. Sinha, and J. C. Wismath, Reduction of the order of dynamic models, *in* "Proc. 1970 National Elec. Conf.," Univ. of Waterloo, 1970.

6. M. Aoki, Control of large scale dynamic systems by aggregation, *IEEE Trans. Automatic Control* **AC-13** (1968), 246–253.
7. R. Bellman, "Introduction to Matrix Analysis," 2nd ed., McGraw–Hill, New York, 1970.
8. H. Neudecker, Some theorems on matrix differentiation with special reference to Kronecker matrix products, *J. Amer. Statist. Assoc.* **64** (1969), 953–963.
9. H. Neudecker, A note on Kronecker matrix products and matrix equation systems, *SIAM J. Appl. Math.* **17** (1969), 603–607.
10. P. Nicolaides and T. T. Soong, Selective subsystem analysis of large transportation systems, submitted for publication.